Neural network enhanced Bayesian global analysis of EKRT+viscous hydrodynamics model of relativistic heavy ion collisions

Jussi Auvinen

University of Wrocław

In collaboration with Henry Hirvonen (Vanderbilt U.)
and Kari J. Eskola and Harri Niemi (U. of Jyväskylä)

December 14, 2025

Uniwersytet Wrocławski

POLISH NATIONAL AGENCY FOR ACADEMIC EXCHANGE

Bayesian analysis

$$\text{Model parameters (input): } \vec{x} = (x_1, ..., x_n)$$
$$\Downarrow$$
$$\text{Model output } \vec{y} = (y_1, ..., y_m) \Leftrightarrow \text{Experimental values } \vec{y}^{\,\text{exp}}$$

Parameter posterior probability distributions from Bayes' theorem:

$$P(\vec{x}|\vec{y}^{\,\text{exp}}) \propto \mathcal{L}(\vec{y}^{\,\text{exp}}, \vec{y}(\vec{x})) \cdot p(\vec{x})$$

- $p(\vec{x})$: Prior probability distribution of $\vec{x}$ (initial constraints)
- $\mathcal{L}(\vec{y}^{\,\text{exp}}, \vec{y}(\vec{x}))$: Likelihood of observation $\vec{y}^{\,\text{exp}}$ for a given $\vec{x}$ (data fit)
- $P(\vec{x}|\vec{y}^{\,\text{exp}})$: Posterior probability of $\vec{x}$ given observation $\vec{y}^{\,\text{exp}}$ (updated belief on plausible values of $\vec{x}$ after observation $\vec{y}^{\,\text{exp}}$)

First task: Define $\vec{x}$, prior probability $p(\vec{x})$, and $\vec{y}$

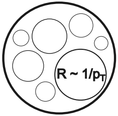
EbyE-EKRT + (2+1)-d viscous hydrodynamics model

H. Hirvonen, K. J. Eskola, H. Niemi, Phys. Rev. C 106, 044913 (2022)

Initial energy density from the event-by-event EKRT minijet saturation model

Niemi et al., Phys. Rev. C 93, 024907 (2016)

- Collinearly factorized NLO pQCD computation of E_T production controlled by cutoff scale $p_{\rm sat}$ determined from the local saturation condition



$$\boxed{\frac{dE_T}{d^2\vec{r}}(p_{\rm sat}, T_A T_A, \sqrt{s_{NN}}, \Delta y) = \frac{K_{\rm sat}}{\pi} p_{\rm sat}^3 \Delta y}$$

Nuclear thickness functions are computed by summing individual nucleon thickness functions: $T_A(\bar{r}) = \sum_i^A T_{n,i}(\bar{r}_i - \bar{r})$, where T_n is a 2-d Gaussian with a width parameter σ_n

- The local energy density $e(\bar{r})$ at formation time $\tau_s(\bar{r}) = 1/p_{\rm sat}(\bar{r}, K_{\rm sat})$ is

$$\boxed{e(\bar{r}, \tau_s(\bar{r})) = \frac{K_{\rm sat}}{\pi}[p_{\rm sat}(\bar{r}, K_{\rm sat})]^4}$$

which is further evolved to a global hydro initialization time $\tau_0 = 1/p_{\rm sat,min} \approx 0.2$ fm using (0+1)-d Bjorken expansion

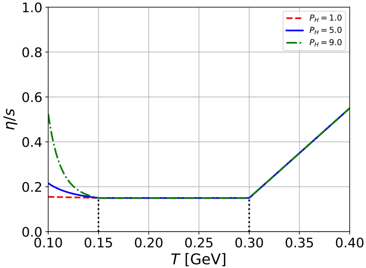
EbyE-EKRT + (2+1)-d viscous hydrodynamics model

We're particularly interested in the temperature dependence of shear viscosity η/s and bulk viscosity ζ/s. Our parametrisation of η/s has 6 parameters:

$$(\eta/s)(T) = \begin{cases} (\eta/s)_{\min} + S_H T \left(\left(\frac{T}{T_H}\right)^{-p_H} - 1 \right), & T < T_H \\ (\eta/s)_{\min}, & T_H \leq T \leq T_H + W_{\eta\,\min} \\ (\eta/s)_{\min} + S_Q(T - (T_H + W_{\eta\,\min})), & T > T_H + W_{\eta\,\min} \end{cases}$$

Example:

- $(\eta/s)_{\min} = 0.15$
- $T_H = 150$ MeV
- $W_{\eta\,\min} = 150$ MeV
- $S_H = 0.1$ GeV^{-1}
- $S_Q = 4.0$ GeV^{-1}



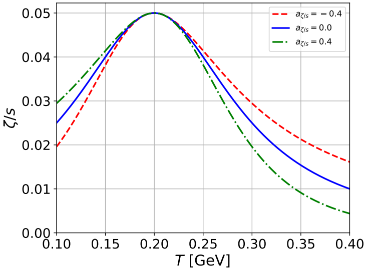
EbyE-EKRT + (2+1)-d viscous hydrodynamics model

Temperature dependence of bulk viscosity has 4 parameters:

$$(\zeta/s)(T) = \frac{(\zeta/s)_{\max}}{1 + \left(\frac{T - T_{\zeta\,\max}}{w(T)}\right)^2}, \quad w(T) = \frac{2(\zeta/s)_{\rm width}}{1 + \exp\left(\frac{a_{\zeta/s}(T - T_{\zeta\,\max})}{(\zeta/s)_{\rm width}}\right)}$$

Example:

- $(\zeta/s)_{\max} = 0.05$
- $T_{\zeta\,\max} = 200$ MeV
- $(\zeta/s)_{\rm width} = 100$ MeV



EbyE-EKRT + (2+1)-d viscous hydrodynamics model

The fluid dynamical equations include local conservation equations for energy, momentum and conserved charges, and evolution equations for the dissipative effects (shear and bulk viscosity). In addition, an equation of state is needed which describes the transition from partonic matter to hadronic matter.

For EoS we use the **s95p** parametrization with a variable chemical freeze-out temperature T_{chem}

P. Huovinen and P. Petreczky, Nucl. Phys. A **837**, 26 (2010)

Dynamical freeze-out conditions: Assuming the mean free path is proportional to the relaxation time τ_π, we have

- Knudsen number condition (local): $\mathrm{Kn} = \tau_\pi \nabla_\mu u^\mu = C_{\mathrm{Kn}}$
- System size condition (non-local): $\frac{\gamma \tau_\pi}{R} = C_R$, $R = \sqrt{\frac{A_{\mathrm{Kn}}}{\pi}}$, where A_{Kn} denotes the area in (x, y)-plane where $\mathrm{Kn} < C_{\mathrm{Kn}}$ (there may be multiple separate areas)

The input

In total, we have 15 free parameters in the model:

- Initial state (2 parameters): $K_{\mathrm{sat}}, \sigma_n$
- Temperature dependence of shear viscosity (6 parameters): $(\eta/s)_{\mathrm{min}}, T_H, S_H, p_H, W_{\eta\,\mathrm{min}}, S_Q$
- Temperature dependence of bulk viscosity (4 parameters): $(\zeta/s)_{\mathrm{max}}, T_{\zeta\,\mathrm{max}}, (\zeta/s)_{\mathrm{width}}, a_{\zeta/s}$
- Chemical freeze-out temperature T_{chem}
- Dynamical freeze-out conditions (2 parameters): C_{Kn}, C_R

The output

Four collision systems:
Au+Au at 200 GeV, Pb+Pb at 2.76 TeV and 5.02 TeV, and Xe+Xe at 5.44 TeV

9 centrality classes: $(0-5)\%, (5-10)\%, (10-20)\%, (20-30)\%, ..., (70-80)\%$

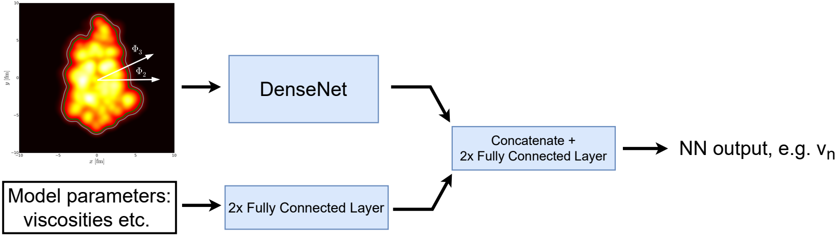
We include the following observables:

- Charged particle multiplicity $dN_{\rm ch}/d\eta$ and $\langle p_T \rangle_{\rm ch}$
- Identified particle multiplicities dN/dy of pions (π^+), kaons (K^+) and protons (p)
- Average transverse momenta $\langle p_T \rangle$ of pions (π^+), kaons (K^+) and protons (p)
- p_T-averaged charged particle 2-particle cumulant flow harmonics, $v_2\{2\}$, $v_3\{2\}$, and $v_4\{2\}$
- Normalized symmetric cumulant $NSC(2,4) = \frac{\langle v_2^2 v_4^2 \rangle - \langle v_2^2 \rangle \langle v_4^2 \rangle}{\langle v_2^2 \rangle \langle v_4^2 \rangle}$

Neural network surrogate model

The flow observables $v_n\{2\}$, $NSC(2,4)$ require a lot of events to constrain statistical errors, which makes them computationally expensive

- Use neural networks to obtain quick estimates for single-event model output for any given energy density profile and parameter combination



H. Hirvonen, K. J. Eskola, H. Niemi, EPJ Web Conf. 296, 02002 (2024)

- For training, sample 1050 points from the input parameter space and run full simulations (40 events / training point for each collision energy)
- Full simulation: ~ 2 events / CPU-hour vs. neural network: ~ 100 events / s on a GPU (NVIDIA V100)

Posterior probability from Bayes' theorem:

$$P(\vec{x}|\vec{y}^{\,\text{exp}}) \propto \mathcal{L}(\vec{y}^{\,\text{exp}}|\vec{x})p(\vec{x})$$

Prior probability $p(\vec{x})$: Uniform hypercube

Likelihood function:

$$\mathcal{L}(\vec{y}^{\,\text{exp}}|\vec{x}) = \frac{1}{\sqrt{|2\pi\Sigma|}} \exp\left(-\frac{1}{2}(\vec{y}(\vec{x}) - \vec{y}^{\,\text{exp}})^T \Sigma^{-1} (\vec{y}(\vec{x}) - \vec{y}^{\,\text{exp}})\right)$$

where Σ is the covariance matrix containing the experimental and theoretical uncertainties (including model emulation uncertainties)
We include theory uncertainty which is (10/20/30)% of the data $\vec{y}^{\,\text{exp}}$

The total likelihood is a product of the likelihoods calculated at each collision energy:

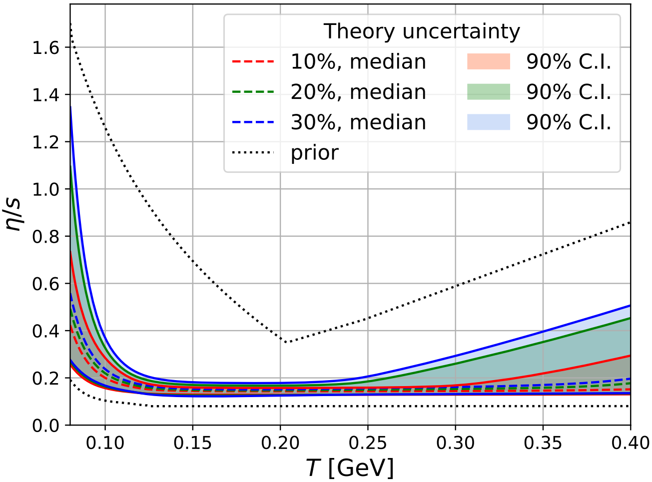
$$\mathcal{L}_{\text{tot}} = \mathcal{L}_{200} \times \mathcal{L}_{2760} \times \mathcal{L}_{5020} \times \mathcal{L}_{5440}$$

Use Markov chain Monte Carlo combined with Gaussian process model emulators to draw samples from the posterior probability distribution

Neural networks: Single-event predictions
for a given energy density profile and parameter values
Gaussian processes: Event-average predictions for given parameter values

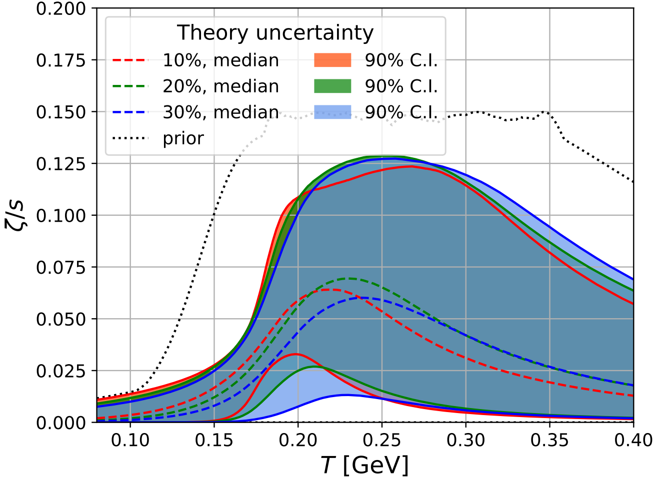
Shear viscosity from posterior distribution

Constant-value plateau between $\approx (150 - 200)$ MeV with $0.12 < (\eta/s)_{\min} < 0.18$; likely to remain near constant also at higher temperatures

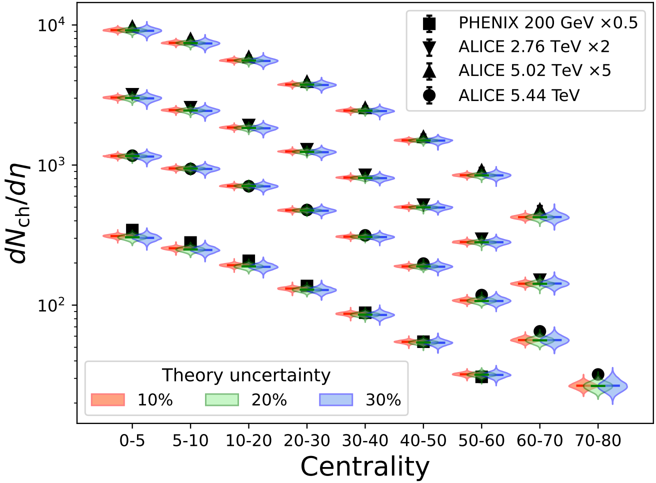


Bulk viscosity from posterior distribution

No strong constraints; non-zero in the temperature range $T \approx (200 - 300)$ MeV



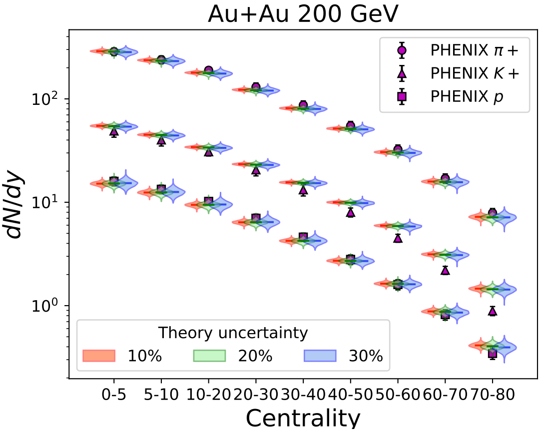
Charged particle yields vs. centrality



PHENIX data: PRC 71, 034908 (2005)

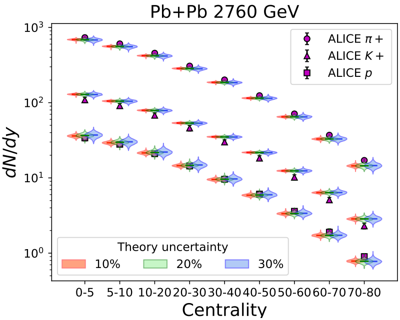
ALICE data: [2.76] PRL 106, 032301 (2011); [5.02] PRL 116, 222302 (2016); [5.44] PLB 788, 166 (2019)

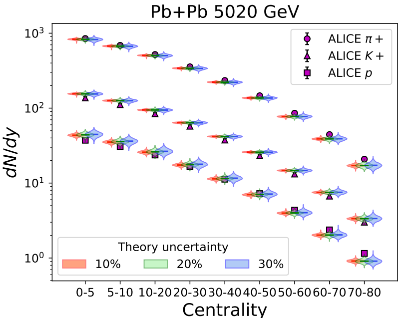
Identified particle yields vs. centrality



PHENIX data: Phys.Rev.C 69, 034909 (2004)

Identified particle yields vs. centrality





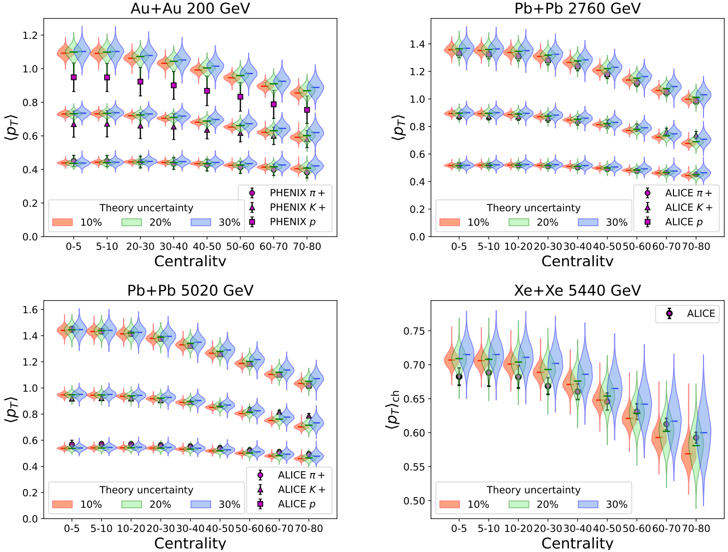
ALICE data: Phys.Rev.C 88, 044910 (2013)

ALICE data: Phys. Rev. C 101, no.4, 044907 (2020)

Mean p_T vs. centrality

PHENIX data: Phys.Rev.C 69, 034909 (2004)

ALICE data: Phys.Rev.C 88, 044910 (2013)



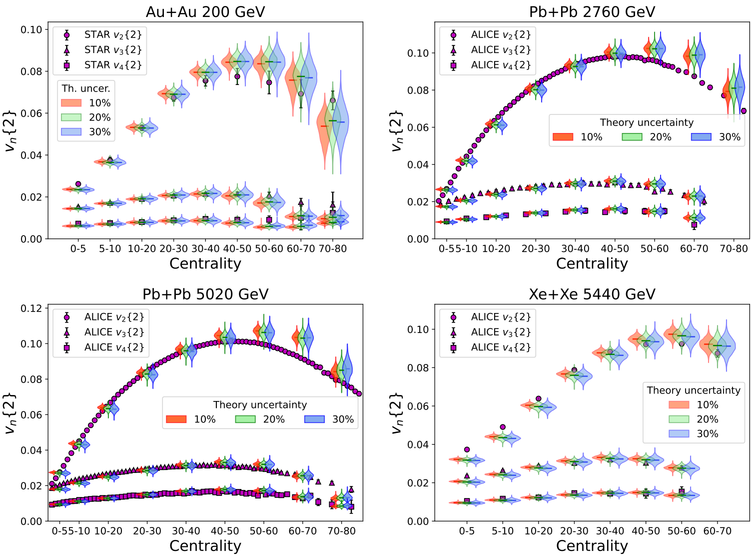
ALICE data: Phys. Rev. C 101, no.4, 044907 (2020)

ALICE data: Phys.Lett.B 788, 166 (2019)

Flow v_n vs. centrality

STAR data: Phys.Rev.C 98, 034918 (2018)

ALICE data: JHEP 07, 103, (2018)

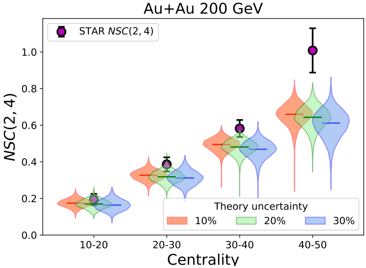


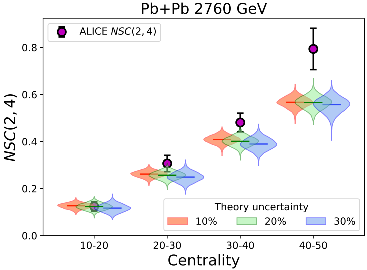
ALICE data: JHEP 07, 103, (2018)

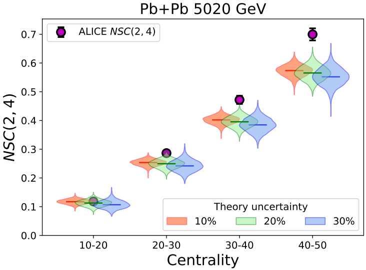
ALICE data: Phys.Lett.B 784, 82 (2018)

$NSC(2,4)$ vs. centrality

STAR data: Phys.Lett.B 783, 459 (2018)







ALICE data: Phys.Rev.Lett. 117, 182301 (2016)

ALICE data: Phys.Lett.B 818, 136354 (2021)

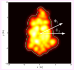
Summary: NN-enhanced Bayesian analysis workflow

Neural networks

(1) Sample training points
from the 15-D input parameter space
(EbyE-EKRT IC w 2+1 D hydro)

(2) Run full event-by-event
simulations for each training point

Single-event energy density profile →



→ DenseNet →

Model parameters in one training point → Model parameters: viscosities etc. → 2x Fully Connected Layer →

Concatenate + 2x Fully Connected Layer → NN output, e.g. v_n

(3) Train neural networks to predict **single-event** output at training points

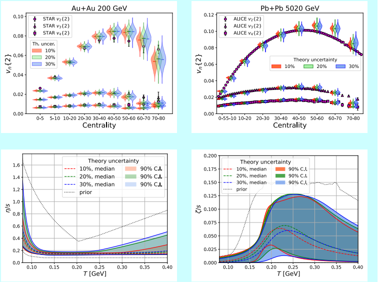
(4) To increase statistics, predict model output for additional initial state profiles using neural networks (can generate ~100 events / s using GPUs)

Bayesian analysis

(5) Based on the ensemble of single-event NN predictions, calculate event-averaged observables for Gaussian process emulator training

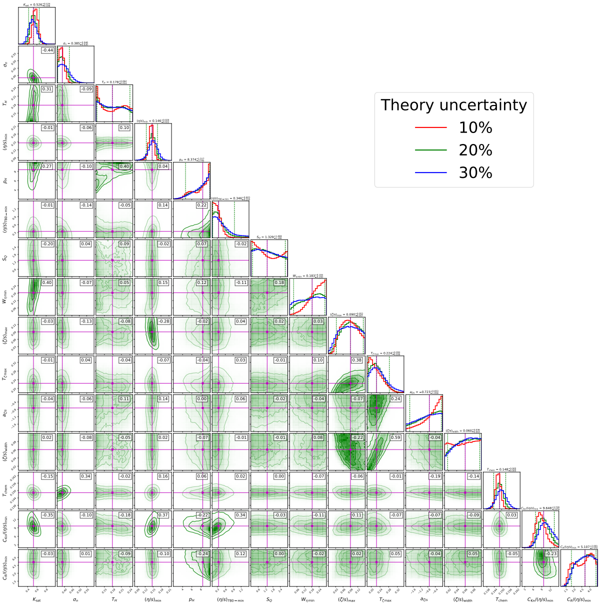
(6) Train Gaussian process emulators
to predict the **event-averaged** output

(7) Run Markov chain Monte Carlo to obtain samples of the posterior distribution, using GP emulators
to evaluate the likelihood at each step

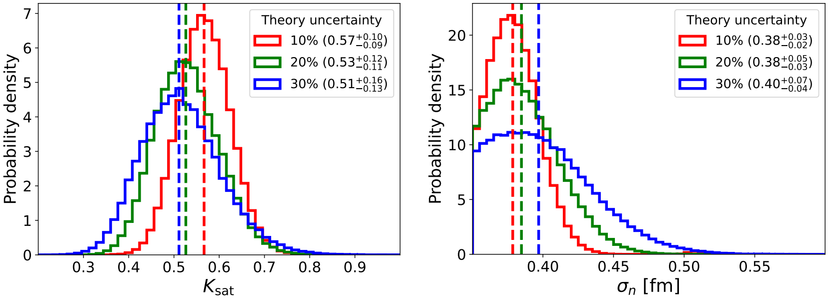


Extra slides

Full posterior distribution



Posterior distributions for initial state parameters



Posterior distributions for freeze-out parameters